

OpenVMS Cluster update

Shyam Sankar G
OpenVMS Engineering



Agenda

- Introduction to OpenVMS Clustering Technology
- OpenVMS Clusters before 8.4
- Challenges with Long Distance OpenVMS clusters
- Addressing challenges with IPCI
- OpenVMS Cluster Technology Overview
- Enhancements in 8.4



OpenVMS Clusters Today



OpenVMS Clusters Today

- SCA (aka SCS) – System Communication Architecture
 - Cluster communication protocol
- Cluster Interconnect
 - Alpha : LAN, Memory Channel, Shared Memory, CI
 - IPF (Integrity) :LAN



OpenVMS Clusters Prior to IPCI

- LAN interconnect for long distance cluster communication
- LAN Bridging and Extended LAN techniques for multi-site long distance clusters
- Nodes belong to same LAN/VLAN for cluster communications

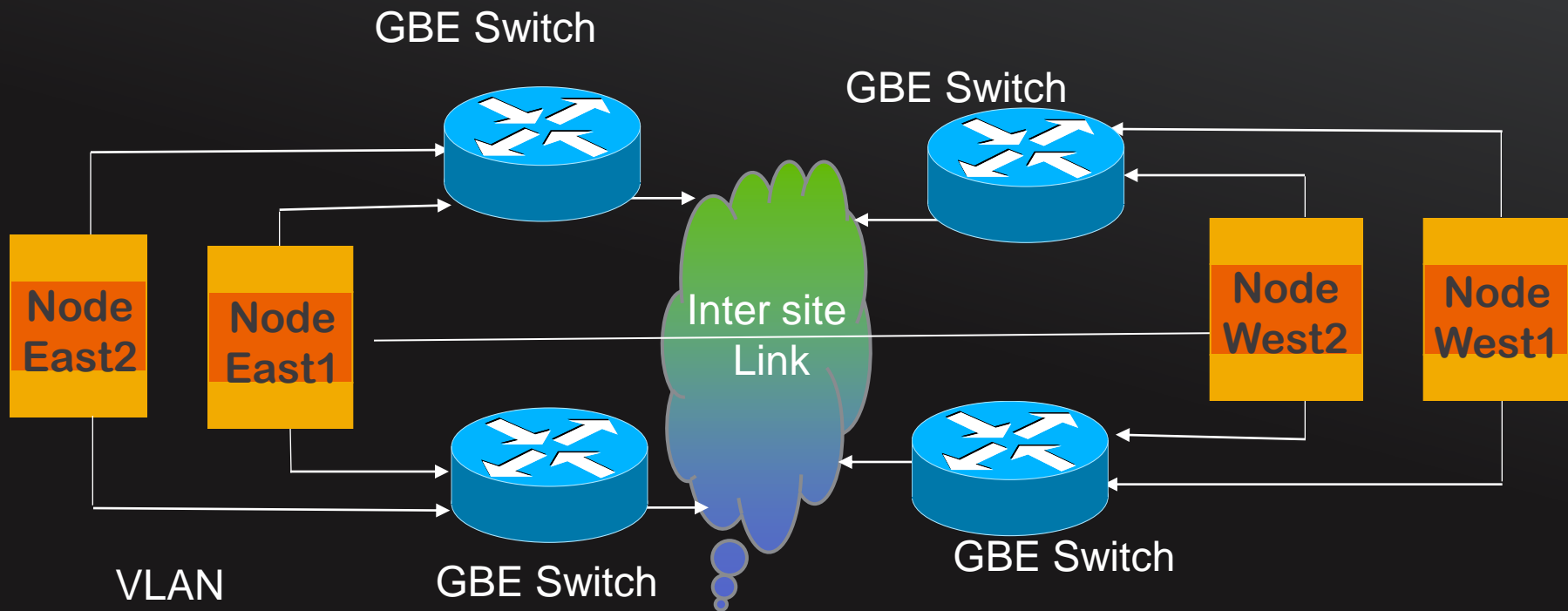


Inter-site Options

- Supported inter site distance can be 500 mile
 - Customer have site separated by 3000 mile
- OpenVMS aggressive failover settings can enable failover in less than 10 seconds
- Inter-site can be linked by:
 - DS-3/T3 (E3 in Europe) or ATM circuits from a telecommunications vendor
 - Microwave link: DS-3/T3 or Ethernet
 - Dark Fiber where available
 - WDM in CWDM or DWDM flavors
- OpenVMS cluster requires 10 MB minimum rate



Disaster Tolerant /Long Distance OpenVMS Clusters



LAN bridging/Extended LANs using switches

Nodes East1, East2, West1, West2 belong to same VLAN

Challenges with OpenVMS Cluster

–Technology

- SCS traffic is Non-IP
- Network switches during higher loads give priority to IP traffic than cluster (SCS) traffic
- Cluster instability during periods of heavy IP usage
- High router utilization for transporting cluster packets
- Corporate policies restricting scope of non-IP protocols



Challenges with OpenVMS Cluster

–Costs

- Extra Cost/license for LAN bridging /layer 2 service
- Specialized hardware and human resource costs for setting up multi site DT cluster with LAN bridging

–Non-availability of LAN bridging from all switch and telco vendors



Addressing Challenges with IPCI

- IPCI is the ability to make use of IP for OpenVMS clusters communications
- IPCI coexists with LAN interconnect for Cluster communication
- IP unicast and optionally IP multicast (administratively scoped) for node discovery
- File based mechanism for unicast node discovery
- No application changes needed



IPCI Benefits

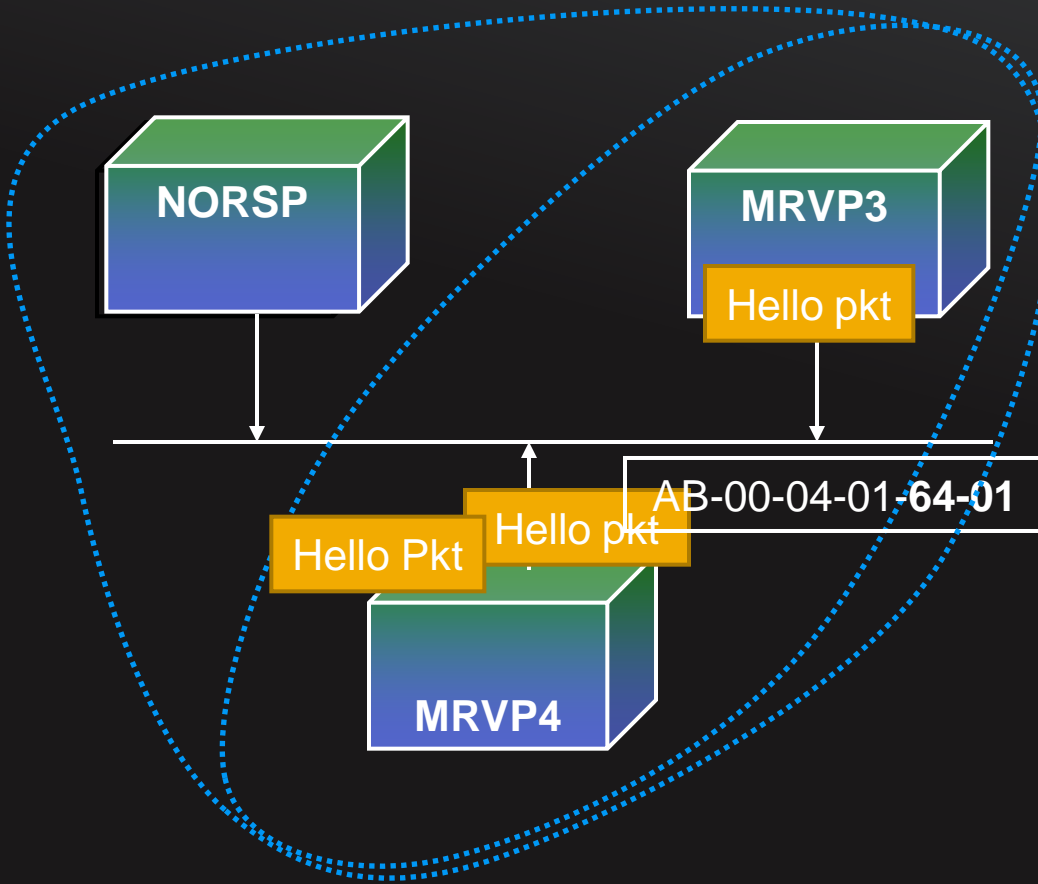
- No more dependence on Vendors with support for LAN bridging
- Lower infrastructural and operational costs
- No extra license/cost for LAN bridging (Layer 2 service)
- Leverage the benefits from the improvements in IP and LAN interconnect technology
- Ability to co-exist in a modern datacentre without special setting



OpenVMS Cluster Technology Overview



OpenVMS Cluster Communication in LAN



MRVP3 and MRVP4 are in cluster

Cluster group number 100

The multicast address is

AB-00-04-01-00-01 +
the cluster group number

AB-00-04-01-**64-01**

Configure NORSP as cluster and
reboot NORSP

Node Discovery – Unicast and Multicast

- IP unicast used for node discovery and hello packets.
- IP multicast can also be used (Administratively scoped IP multicast address)
- Remote nodes not in IP multicast domain use IP unicast technique to join Cluster and send hello packets

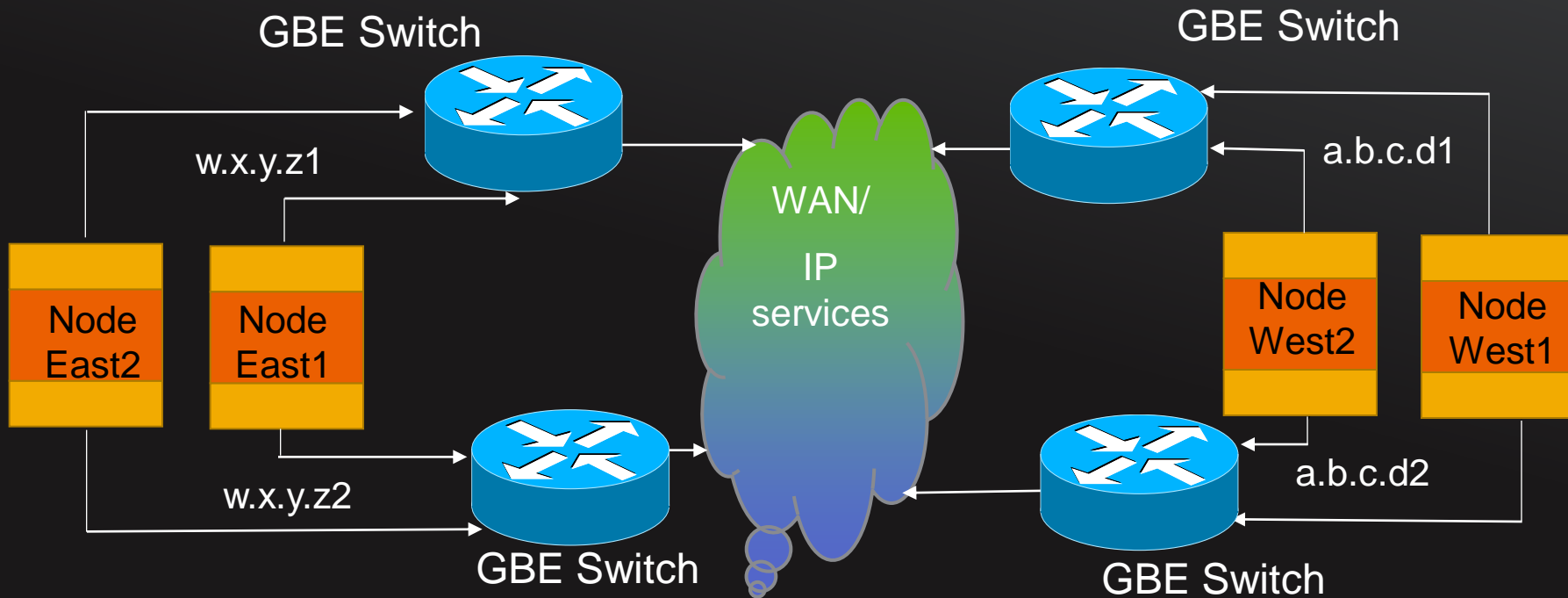


TCP/IP Services boot time loading and Initialization

- TCP/IP services loaded early during boot to facilitate
 - Cluster communications in an IP only network environment
 - Cluster formation in a IP only network
- Ability to make use of boot time configuration information to initialize TCP/IP services
- Existing boot sequence – LAN,PE driver, TCP/IP
- Boot Sequence with IPCI – LAN,TCP/IP, PE driver



Cluster using IPCI



- Node East1, East2, West1, West2 can be part of the same or **different LAN** for cluster communications using IPCI.
- East1 and West 2 has a Virtual Circuit (VC) VC consists of IP channels for SCS traffic

The IPCI Solution

- Requires HP TCP/IP services for OpenVMS V5.7
 - Not available with other TCP/IP stacks at this time
 - Initial release supports IPv4 only; no IPv6
 - Requires static IP addresses and IP Unicast, optionally uses IP Multicast
- Coexists with LAN interconnect for Cluster communication
- Support for Satellite nodes included
- Existing intra-node distance/latency limitation applies
- LAN channels preferred over IP channels

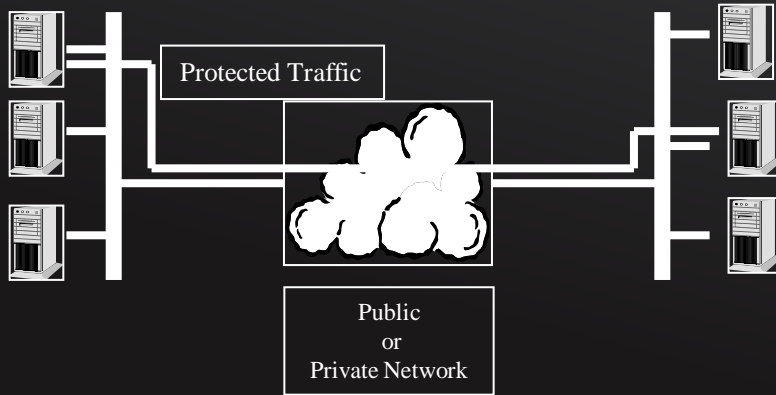


Security Considerations

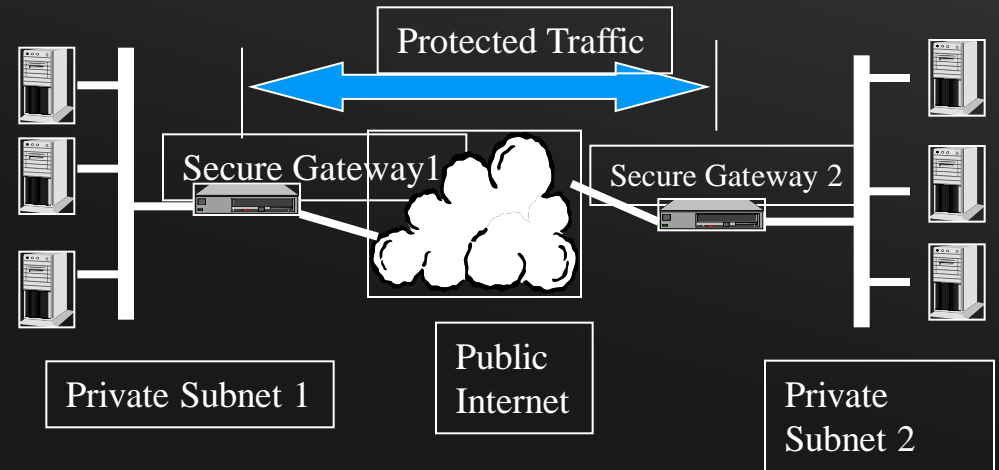
- Normal intranet and Internet Security principles
- VPN (virtual Private Network)
- TTL (Time to live)
- Firewalls



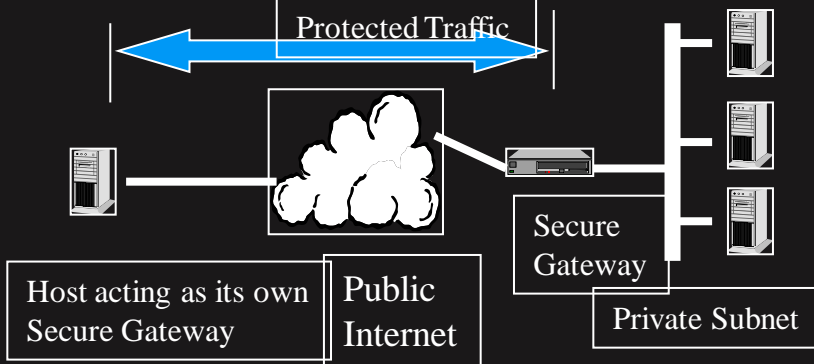
IPsec Security – Connection Methods



IPsec for Host-to-Host



IPsec for Virtual Private Networks



IPsec for Remote Access



Performance

- Engineering has conducted some performance test to recommend configurations for optimal performance
- Observation to date show TCP/IP ping latency close to latency reported by PEdriver
- Try to affinitize the LAN, TCP/IP and PE device on the same CPU
- Increase transmit window
 - (\$MC SCACP SET VC/WINDOW)



PEdriver improvements V8.4!

- ECS stability
- Multiple channels actively used (High Bandwidth)
- Performance benefit of ~50%
- Avoid CLUEXIT crashes with multiple channels
- Available in V8.4 and in TIMA kits for prior versions



PEdriver Performance Improvements V8.4

- Equivalent Channel Set
 - TUK4S:56: \$mc scacp show chan tuk8s
 - TUK4S PEA0 Channel Summary 1-SEP-2010 10:22:07.84:

Remote	Device	Channel	Total	ECS	Priority	Buffer	Delay	Load
TUK8S	EWA EWK	Open	29	Y (T, P, F)	0 0	2 1432	250.0	1000
TUK8S	EWA EWM	Open	28	Y (T, P, F)	0 0	2 1432	250.0	1000



PEdriver Performance Improvements – V8.4

TUK4S:56: \$mc scacp show chan tuk8s/ecs/3

TUK4S PEA0 Channel Equivalent Channel Set (ECS) 1-SEP-2010
10:26:30.99:

Remote Node	Device Loc	ECS Rmt	ECS State	ECS Losses	ECS Transitions	Average RTTime	Remote Ring Size	Remote Rcv Cache	Time in Open State	ECS Member Time
TUK8S	EWA	EWK	Y(T,P,F)	0	5	250.0	255	64	13:55:14.41	13:51:55.77
TUK8S	EWA	EWM	Y(T,P,F)	0	5	250.0	255	64	13:55:14.41	13:51:55.77



PEdriver improvements V8.4

– \$ mc scap show chan MIKUMI/ecs/3 (OpenVMS 8.3)

NAMIB PEA0 Channel Equivalent Channel Set (ECS) 1-SEP-2010 10:30:05.34:

Remote Node Time	LAN Loc	Dev Rmt	ECS State	ECS Losses	ECS Transitions	Average RTTime	Remote Ring Size	Remote Rcv Cache	Rsvp Thresh	Channel Open Time	ECS Member
MIKUMI	EIA	EWA	Y(T,P,F)	0	3793	6468.0	16	32	0	22:49:39.82	20:23:47.14
MIKUMI	EIB	EWB	Y(T,P,F)	0	3643	6567.5	16	32	0	22:49:45.67	20:48:26.37



ICC SDA Extension

Available in v8.4

SDA> icc

ICC SDA Extension - Quick Help Information

ICC SDA commands:

- ICC SHOW PROCESSES - displays processes using ICC
- ICC SHOW ASSOCIATIONS - displays open ICC associations
- ICC SHOW CONNECTIONS - displays open ICC connections
- ICC HELP - more information about ICC SDA commands

SDA>

SDA> icc show processes

ICC Process Summary

Extended	Indx	Process name	Username	Assoc	Conn	ICCPDB
-- PID --	--	-----	-----	-----	-----	-----
20400433	0033	SYSTEM	SYSTEM	1	1	90612DC0



ICC SDA Extension (contd...)

```
SDA> icc show associations
```

```
ICC Associations
```

```
-----
```

```
--- ICCPAB Summary Page ---
```

ICCPAB Addr	Extended Process name	State	Conn Association Name
-----	---PID---	-----	-----
9054DE80	20400433 SYSTEM	Open	1 ICC\$PID_20400433

```
ICC Associations
```

```
-----
```

```
--- ICC Process Association Block (ICCPAB) 9054DE80 ---
```

```
State: 0001 Open           Association Name: ICC$PID_20400433
EPID: 20400433           Process Name: SYSTEM
Connections: 1           First ICCPCB: 90685B80
ICCPDB: 90612DC0
```

```
Logical Name:
```



ICC SDA Extension (contd...)

SDA> icc show connections

ICC Connections

--- ICCPCB Summary Page ---

ICCPCB Addr	Extended	Process name	State	Node	Association Name
90685B80	20400433	SYSTEM	Open	OOTY	ICC\$PID_20400433

ICC Connections

--- ICC Process Connection Block (ICCPCB) 90685B80 ---

State:	0001 Open	Association Name:	ICC\$PID_20400433
EPID:	20400433	Process Name:	SYSTEM
Remote Node:	OOTY	Remote Association:	ICC\$PID_2020042A
ICCPDB:	90612DC0	ICCPAB:	9054DE80
LinkID:	00000301	LCB:	9069FC00
CDT:	9069FC80	Open for:	0 0:01:04
Local ConID:	C4A3000B	Remote ConID:	C4A3000B
Bytes Sent:	10074	Messages Sent:	69
Bytes Rcvd:	10074	Messages Rcvd:	69

SDA>



Q&A

